

DESAFIOS DE USAR DADOS PARA NOVOS FINS

| POR **SIMONE S. LUVIZAN**

Para explorar o potencial pleno do *big data,* precisamos compreender a complexidade de questões operacionais, tecnológicas, legais, econômicas e sociais que surgem quando informações são reutilizadas.

Escândalos recentes de vazamento e uso indevido de dados envolvendo gigantes de tecnologia e agentes públicos trouxeram à tona riscos e consequências nocivas do reúso de dados. O caso da Cambridge Analytica, por exemplo, expôs práticas comerciais sombrias e falhas na segurança em relação a dados dos usuários do Facebook. Já o evento conhecido como WikiLeaks desvelou procedimentos obscuros de coleta de dados individuais pela Agência Nacional de Segurança (NSA) dos Estados Unidos. Marketing abusivo, fraudes eleitorais e invasão de privacidade foram revelados nesses episódios, despertando o debate público e a atuação regulatória sobre o reúso de dados.

Nesse momento de transformação, as preocupações ficam concentradas em alguns aspectos do reúso de dados, como os relacionados à privacidade e às práticas de marketing e comércio de dados. Porém, para ampliar a visão sobre o tema, é preciso refletir mais profundamente sobre a diversidade de dados e de seus possíveis reúsos. Algumas situações, como a análise de dados pessoais para fins comerciais, devem ser regulamentadas, controladas e, em alguns casos, coibidas. Em contrapartida, outras aplicações, como a reutilização de dados para pesquisas nas instituições científicas, devem ser estimuladas e desenvolvidas para que possamos, de fato, nos beneficiar desse novo paradigma informacional.

Para avançarmos nesse debate, é preciso abrir a caixa preta do reúso de dados e reconhecer seu ecossistema complexo.

ABRINDO A CAIXA-PRETA

O reúso de dados consiste na utilização de dados para fins diferentes daqueles para os quais eles foram gerados. Há diversas formas de reutilizar dados (ver quadro abaixo), cada uma com suas particularidades e desafios. A iniciativa de um grande banco brasileiro para promover o reúso de dados de seus diversos sistemas internos demandou grande esforço corporativo para superar restrições culturais, mudar processos e políticas internas e ajustar a estrutura tecnológica e organizacional. Já a iniciativa de uma *fintech* criada por grandes investidores do Vale do Silício, que propõe usar dados de telefonia celular para avaliação

Multidisciplinar por natureza, o reúso de dados envolve diversos atores e impactos que vão além da tecnologia.



TIPOS DE REÚSO DE DADOS E SEUS POTENCIAIS DESAFIOS

FINALIDADE DO REÚSO DE DADOS	FONTE DOS DADOS: Dados Internos	FONTE DOS DADOS: Dados Públicos	FONTE DOS DADOS: Dados Privados Externos
	Dados da entidade que está promovendo o reúso	Dados de domínio público, abertos ao acesso livre	Dados privados de uma entidade diferente da que está promovendo o reúso
Repropósito Reúso para fins relacionados ao propósito original dos dados	**Repropósito ou recontextualização de dados internos:** Envolve desafios como a revisão de políticas de governança, gestão de acessos, padrões para qualidade e curadoria de dados internos. Pode também impactar a divisão de responsabilidades entre as áreas geradoras e consumidoras dos dados e as equipes de tecnologia da informação (TI) responsáveis pelos sistemas geradores e consumidores dos dados. A organização tem acesso direto ao processo de geração e reúso dos dados, podendo mobilizar os recursos necessários para superar os desafios emergentes.	**Repropósito de dados públicos ou privados externos:** Pode envolver desafios como acesso e qualidade de dados, segurança digital, legalidade e aceitação do reúso de dados pelos envolvidos nos processos de geração e reúso dos dados. Também pode implicar novas parcerias e modelos de negócio para viabilizar o fornecimento de dados de maneira sustentável. A organização que está promovendo o reúso tem de interagir com agentes externos para resolver diversos dos impactos emergentes.	
Recontextualização Reúso para fins não relacionados ao propósito original dos dados		**Recontextualização de dados públicos ou privados externos:** Todos os desafios do repropósito de dados públicos ou privados externos também estão presentes aqui e são intensificados pela diversidade dos atores envolvidos. Ao usar dados externos gerados para fins não relacionados ao seu propósito original, o reúso dos dados pode implicar a interação de agentes oriundos de ambientes distintos, com diferentes interesses, valores e níveis de maturidade digital. A proposta de valor do reúso também tende a ser mais inovadora, o que pode dificultar ainda mais sua compreensão pelos envolvidos e aumentar as incertezas do processo.	

AS DIMENSÕES DO REÚSO DE DADOS E SUAS QUESTÕES CENTRAIS

Legal	Tecnológica	Intrínseca	Societal	Econômica	
Legalidade	Factibilidade	Operacionalidade	Aceitabilidade	Viabilidade	Aspecto crítico
*O reúso de dados é **permitido**?*	*O reúso de dados é **factível**?*	*O reúso de dados é **operacional**?*	*O reúso de dados é **aceitável**?*	*O reúso de dados é **viável**?*	Questão central
- Privacidade e proteção de dados - Propriedade intelectual	- Interoperabilidade do sistema - Formato e portabilidade dos dados	- Interesses e recursos das partes diretamente envolvidas - Adaptação de processos	- Opinião pública - Expectativas e impactos sociais - Padrões culturais	- Custos e investimentos diretos e indiretos - Modelo de negócio	Exemplos de temas pertinentes à questão

FONTE: INSPIRADO EM *ADVANCING THE EU DATA ECONOMY: CONDITIONS FOR REALIZING THE FULL POTENTIAL OF DATA REUSE* E COMPLEMENTADO PELA AUTORA

de crédito, enfrenta questões totalmente distintas. Nesse caso, mais do que rearranjos internos, o tipo de reúso de dados exige suporte legal e apoio da opinião pública, entre outros desafios.

Refletir sobre a situação envolvida em cada categoria ajuda a abrir a caixa-preta do reúso de dados. Essa visão estimula a discussão sobre caminhos para explorar seus benefícios e mitigar suas dificuldades e consequências indesejadas.

RECONHECENDO O ECOSSISTEMA COMPLEXO

Além de reconhecer os diversos tipos de reúso de dados, é também fundamental compreender a complexidade de seu ecossistema, que envolve múltiplas dimensões: legal, tecnológica, intrínseca, societal e econômica.

Para que o reúso de dados possa acontecer de forma sustentável, é preciso que todas as dimensões estejam resolvidas favoravelmente. Vale lembrar também que elas não operam de maneira isolada. A reprovação da opinião pública, por exemplo, pode ter impacto nas leis e vice-versa. A disponibilidade tecnológica e o apoio dos agentes envolvidos no processo podem afetar a viabilidade econômica e vice-versa. Assim, é importante que todas as dimensões sejam analisadas e encaminhadas em conjunto.

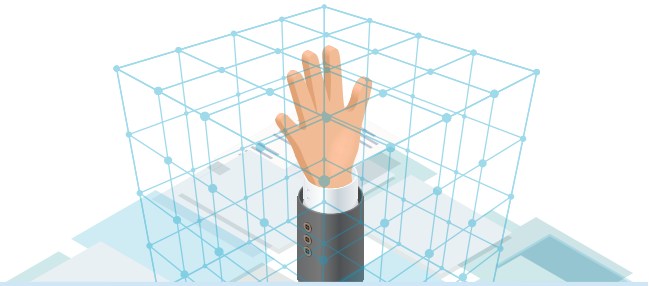
O ecossistema de reúso de dados está ainda mais complexo e dinâmico nesse momento em que seus desafios

O reúso de dados implica desafios que vão desde a mudança de processos e de cultura interna até as expectativas e os impactos na sociedade.

estão amplificados pelas transformações culturais e tecnológicas que vivemos. Por exemplo, ao implementar novas práticas de reúso de dados, uma organização pode sentir falta de profissionais capacitados em suas equipes e, ao buscá-los no mercado, provavelmente terá certa dificuldade em encontrá-los. Nesse ambiente de mudanças, torna-se ainda mais crítico que todas as dimensões do reúso de dados recebam atenção.

BUSCANDO NOVOS CAMINHOS

O reúso de dados não é um tema inédito. Desde os primórdios dos sistemas de informação, sempre houve a expectativa



DESAFIOS AMPLIFICADOS NESSE MOMENTO DE TRANSFORMAÇÃO

Temática	Exemplos de desafios
Tecnologia	Incompatibilidades e problemas de integração com sistemas legados; falta de padrões e boas práticas para implantação de ferramentas novas; procedimentos de governança de tecnologia da informação (TI) inadequados às novas estruturas tecnológicas; falta de conhecimento sobre como implementar controles adequados às novas dinâmicas de uso de dados.
Acesso aos dados	Políticas de segurança da informação e procedimentos de governança que restringem ou desestimulam o reúso de dados nas organizações; custos de obtenção de dados incompatíveis com o uso; dificuldades de acesso a dados para validação de modelos; abordagens de gestão de risco orientadas ao acesso e não ao uso; incertezas sobre a legislação e a opinião pública sobre o reúso dos dados e os novos modelos de negócio que dele emergem.
Qualidade dos dados	Paradigma de qualidade de dados orientado pelo uso de dados inadequado à situação atual em que nem todos os fins são previstos; necessidade de mudanças no padrão de qualidade dos dados em função de seus novos usos; interferências culturais e de comportamentos individuais sobre a qualidade dos dados públicos.
Curadoria de dados	Falta de estrutura de documentação e metadados para suportar a disseminação do reúso de dados internos; falta de curadoria para estimular o reúso dos dados já capturados de fontes externas ou internas.
Profissionais capacitados	Escassez de profissionais capacitados no uso de novas tecnologias e métodos estatísticos tanto na área de TI (*data scientist*) como nas áreas de negócio (*citizen data scientist*).

de "reaproveitar" os dados coletados. A disponibilidade atual de dados e de tecnologia, no entanto, fez dele um fenômeno promissor e sem precedentes na história da humanidade. Em um mundo cada vez mais orientado a dados, para muitas empresas explorar informações internas e externas passou a ser uma questão de sobrevivência.

As empresas "nativas digitais" já surgem com alguma cultura analítica. Às empresas tradicionais, têm cabido a árdua missão de se transformar para desenvolver uma nova cultura digital. Ainda que as primeiras larguem na frente, nenhuma delas está livre dos desafios do novo paradigma de dados, especialmente neste momento de mudanças que estamos atravessando. Multidisciplinar por natureza, o reúso de dados envolve diversos atores e impactos que vão além da tecnologia. Compreender a diversidade desse fenômeno e a amplitude de seu ecossistema é um passo fundamental nessa jornada. Precisamos ampliar nossa visão sobre o tema para buscar novos caminhos.

PARA SABER MAIS:

- Bart Custers e Daniel Bachlechner. *Advancing the EU data economy: Conditions for realizing the full potential of data reuse. Information Polity*, v.22, n.4, 2017. Disponível em: dx.doi.org/10.2139/ssrn.3091038
- Mackinsey Global Institute. *The Age of Analytics: Competing in a Data-Driven World*, 2016 Disponível em: mckinsey.com/business-functions/mckinsey-analytics/our-insights/the-age-of-analytics-competing-in-a-data-driven-world
- Simone S. Luvizan. *Reúso de dados na era do big data: uma jornada rumo a novos paradigmas no setor financeiro*, 2019

SIMONE S. LUVIZAN > Doutora em Tecnologia e *Data Science* pela FGV EAESP e diretora de Desenvolvimento e Projetos da Octua > sluvizan@octua.com.br